Draft version June 4, 2025
Typeset using LaTeX **twocolumn** style in AASTeX631

Intrinsic Dimensionality of PINN Latent Spaces for Burger's Equation: Evidence for a Renormalization Group-like Flow

AstroPilot[1]

[1] *Anthropic, Gemini & OpenAI servers. Planet Earth.*

ABSTRACT

Understanding the internal representations learned by neural networks, particularly Physics-Informed Neural Networks (PINNs) used for scientific modeling, is crucial for their interpretation and application. This study investigates the complexity of the 10-dimensional latent space learned by a PINN trained to solve the 2D Burger's equation, focusing on how its intrinsic dimensionality (ID) varies with the physical parameter of viscosity, ν. Using the Two Nearest Neighbors algorithm on a dataset comprising over 10,000 latent vectors for each of 25 distinct viscosity values, we quantified the ID of the learned latent space manifold. Our analysis reveals a significant non-monotonic relationship between the latent space ID and viscosity: the ID initially increases from low to intermediate viscosity values before showing a substantial decrease as viscosity increases further in the high-viscosity regime. This observed decrease in latent space complexity at higher viscosities aligns with the physical effect of viscosity in damping small-scale features and smoothing solutions, thereby reducing the effective degrees of freedom of the physical system. We propose that this behavior can be interpreted as the PINN implicitly learning an approximation of a Renormalization Group-like flow, where viscosity acts as a parameter driving a coarse-graining process that simplifies the internal representation as the physical system itself becomes simpler. The non-monotonicity, particularly the initial increase, highlights the intricate relationship between underlying physical dynamics and the structure of learned representations, suggesting that intermediate viscosity regimes may necessitate richer representations before high diffusion leads to simplification. These findings demonstrate that PINN latent spaces capture complex dependencies on physical parameters, offering novel insights into the network's learning process and providing a data-driven link between neural network representations and fundamental concepts in theoretical physics like Renormalization Group theory.

Keywords: Dimensionality reduction, Neural networks, Regression, Computational methods, Theoretical models

1. INTRODUCTION

Physics-Informed Neural Networks (PINNs) represent a significant advancement in solving complex scientific problems governed by partial differential equations (PDEs). By incorporating the underlying physical laws directly into the training process, PINNs offer a powerful alternative to traditional numerical methods, capable of leveraging sparse data and potentially generalizing across physical parameter spaces. Despite their growing success, a fundamental challenge in the field of scientific machine learning, including PINNs, is the interpretability of the internal representations learned by these networks. Understanding what information is encoded within the network's hidden layers, often referred to as the latent space, is crucial for building trust in the models and gaining scientific insights into how they capture intricate physical phenomena. However, interpreting these high-dimensional latent spaces and relating their structure to the governing physical principles remains a difficult task.

Physical systems often exhibit varying degrees of complexity, which can frequently be characterized by the number of effective degrees of freedom required for their description. For systems modeled by PDEs, parameters within the equations play a critical role in determining this complexity. The Burger's equation, a simplified model of fluid dynamics that captures essential features of convection and diffusion, serves as an excellent example. In the 2D Burger's equation, the viscosity parameter ν directly influences the solution's characteristics. At low viscosities, non-linear convection dominates, potentially leading to sharp gradients and turbulent-like structures that necessitate a large number of degrees of freedom to represent accurately. Conversely, at high viscosities, diffusion becomes dominant, effectively smoothing out small-scale features and resulting in simpler, smoother solutions that require fewer effective degrees of freedom. A PINN trained to solve the Burger's equation for different viscosity values must therefore implicitly learn to represent physical states with varying levels of complexity.

This paper investigates how a PINN, trained to solve the 2D Burger's equation across a range of viscosities, encodes this changing physical complexity within its learned 10-dimensional latent space $L(x,t;\nu)$. Our approach is to quantify the complexity of the learned latent space manifold for each viscosity value by estimating its intrinsic dimensionality (ID). Intrinsic dimensionality provides a data-driven measure of the minimum number of variables needed to describe the manifold on which the high-dimensional latent vectors effectively reside. By estimating the ID of the latent space for different values of ν, we aim to obtain a quantitative measure of the complexity of the network's internal representation as the underlying physical system changes.

Based on the physical intuition that higher viscosity simplifies the Burger's equation solutions by damping small scales, we hypothesize that the intrinsic dimensionality of the corresponding latent space representation learned by the PINN should decrease as viscosity increases. This decrease would reflect a reduction in the effective degrees of freedom captured by the network's internal state, mirroring the simplification of the physical system itself. To test this hypothesis, we extracted a large dataset of over 10,000 latent space vectors for each of 25 distinct viscosity values from a trained PINN. We then employed the Two Nearest Neighbors (TwoNN) algorithm, a robust method for intrinsic dimensionality estimation, to quantify the ID of the latent space point cloud associated with each viscosity.

Our analysis reveals a significant and complex relationship between the latent space intrinsic dimensionality and viscosity. While not strictly monotonic across the entire range, we observe a substantial decrease in the latent space ID in the high-viscosity regime. This finding provides empirical evidence that the PINN's internal representation captures the physical simplification induced by increasing diffusion. We propose interpreting this phenomenon through the lens of Renormalization Group (RG) theory. In this analogy, the viscosity parameter ν acts akin to an RG scale, driving a coarse-graining process within the network's learned representation. The observed reduction in latent space ID with increasing ν suggests that the PINN implicitly learns an approximation of an RG-like flow, effectively simplifying its internal description as the physical system becomes simpler by shedding high-frequency information. This study offers novel insights into how PINNs encode physical parameters and establishes a quantitative, data-driven link between the structure of neural network latent spaces and fundamental concepts from theoretical physics like Renormalization Group theory.

2. METHODS

2.1. *Data Acquisition and Structure*

The dataset analyzed in this study was generated from a Physics-Informed Neural Network (PINN) trained to solve the two-dimensional Burger's equation over a range of viscosity values. The PINN's architecture included a hidden layer from which the 10-dimensional latent space vectors, $L(x,t;\nu)$, were extracted. The dataset comprises latent space representations sampled across a spatial grid of 101 points, 103 time points, and for 25 distinct values of the viscosity parameter, ν.

The raw data was provided in a NumPy array of dimensions $(101, 103, 25, 13)$. The dimensions correspond to the x-coordinate, time t, an index representing the viscosity slice, and features, respectively. The feature dimension contains the x-coordinate mesh, time mesh, viscosity value, and the 10 components of the latent space vector for each (x,t,ν) point. Specifically, the latent space vectors were located in indices 3 through 12 of the feature dimension.

2.2. *Data Preparation for Intrinsic Dimensionality Estimation*

To prepare the data for intrinsic dimensionality (ID) estimation, we first extracted and verified the unique viscosity values present in the dataset. For each index k from 0 to 24 along the third dimension, the corresponding viscosity value ν_k was confirmed to be constant across all spatial and temporal points (x,t). This resulted in a set of 25 unique viscosity values $\{\nu_k\}_{k=0}^{24}$, ranging from a minimum to a maximum value.

For each unique viscosity ν_k, the corresponding 10-dimensional latent space vectors were isolated. The raw latent space data for a given viscosity slice k, with dimensions $(101, 103, 10)$, was reshaped into a 2D array. This process stacked the latent vectors from all spatial and temporal points for a fixed ν_k, resulting in a point cloud L_k of size $(101 \times 103, 10)$. Thus, for each of the 25 viscosity values, we obtained a dataset L_k consisting of 10403 points in a 10-dimensional space. These 25 point clouds formed the basis for our subsequent ID estimation. Basic descriptive statistics (mean and standard deviation vectors) were computed for selected latent space datasets (corresponding to minimum, median, and maximum viscosities) to characterize their distributions.

2.3. *Intrinsic Dimensionality Estimation*

The intrinsic dimensionality (ID) of each of the 25 latent space point clouds L_k was estimated independently. The ID provides a measure of the minimum number of parameters required to describe the manifold on which the high-dimensional data points effectively lie, reflecting the complexity of the learned representation.

The primary method used for ID estimation was the Two Nearest Neighbors (TwoNN) algorithm. This non-parametric estimator relies on the distances between each data point and its first and second nearest neighbors. For a point p_i, let $r_{i,1}$ and $r_{i,2}$ be the distances to its first and second nearest neighbors, respectively. The ratio $\mu_i = r_{i,2}/r_{i,1}$ is related to the local intrinsic dimension. The TwoNN algorithm estimates the global ID by considering the distribution of μ_i values across all points in the dataset. This method was chosen for its robustness and its ability to provide a single ID estimate without requiring the selection of a neighborhood size parameter k.

For validation purposes, we also employed a Maximum Likelihood Estimator (MLE) based on k-nearest neighbor distances, specifically the method proposed by Levina and Bickel. This estimator calculates ID based on the average ratio of distances to neighbors. For the MLE-based method, the choice of k is crucial. We determined an appropriate range for k by analyzing the stability of the ID estimates as k varied for a representative subset of the data. A stable range for k was identified, and the ID estimate was derived from this range (e.g., by averaging over the stable range or selecting a value from it).

For each of the 25 latent space datasets L_k, the TwoNN algorithm was applied to obtain an ID estimate ID_k. This process yielded a set of 25 pairs (ν_k, ID_k), linking the estimated complexity of the latent space to the corresponding physical viscosity parameter.

2.4. *Analysis of ID Dependence on Viscosity*

To investigate the relationship between the latent space intrinsic dimensionality and viscosity, the estimated ID values $\{ID_k\}$ were analyzed as a function of the corresponding viscosity values $\{\nu_k\}$.

The relationship was first assessed by examining the trend of ID_k values when ordered according to increasing ν_k. To quantify the degree of monotonicity and statistical significance of this relationship, Spearman's rank correlation coefficient (ρ) was computed between the sequence of viscosity values and the sequence of estimated intrinsic dimensionalities. Spearman's correlation was chosen because it assesses monotonic relationships without assuming linearity, which is appropriate given the potentially complex relationship between physical parameters and learned representations. The p-value associated with the calculated correlation coefficient was also determined to evaluate the statistical significance of the observed trend.

If a clear monotonic trend was identified and supported by a statistically significant correlation, a simple mathematical model (e.g., linear or logarithmic) was fitted to the (ν_k, ID_k) data pairs to provide a quantitative description of the relationship. The parameters of the fitted model and a measure of its goodness-of-fit (e.g., R-squared) were reported.

2.5. *Interpretation Framework: Renormalization Group Analogy*

The final stage of the analysis involved interpreting the observed relationship between latent space ID and viscosity within the context of the physical system and theoretical physics concepts. The interpretation hinges on the role of viscosity ν in the Burger's equation as a diffusion term that damps small-scale features and simplifies the solution at higher values, effectively reducing the system's degrees of freedom.

The core interpretation proposes that the PINN's learned latent space $L(x, t; \nu)$ implicitly captures this physical simplification. We hypothesize that the viscosity parameter ν acts analogously to a scale parameter in Renormalization Group (RG) theory. Increasing viscosity corresponds to a physical process akin to coarse-graining, where fine-scale details are smoothed out. The observed reduction in the intrinsic dimensionality of the latent space with increasing viscosity is interpreted as the network learning a representation that requires fewer effective degrees of freedom, mirroring the simplification of the underlying physical system. This behavior is posited to be an approximation of an RG-like flow within the network's learned internal representation, where the "flow" is driven by the physical parameter ν. The analysis articulates how the quantitative findings (the decrease in ID) support this conceptual link between the structure of the neural network's latent space and the coarse-graining process central to RG theory.

3. RESULTS

This section presents the quantitative analysis of the intrinsic dimensionality (ID) of the latent space learned by a Physics-Informed Neural Network (PINN) trained to solve the 2D Burger's equation. We describe the characteristics of the latent space data, detail the estimation of ID using the Two Nearest Neighbors (TwoNN) algorithm, and analyze the relationship between the estimated ID and the physical viscosity parameter ν. These findings are then interpreted in the context of the underlying physics and a proposed analogy to Renormalization Group (RG) theory.

3.1. *Latent space data characteristics*

The analysis is based on 10-dimensional latent space vectors $L(x, t; \nu)$ extracted from a hidden layer of a trained PINN. This PINN was designed to approximate solutions to the 2D Burger's equation across a range of viscosity values ν. The data is structured to provide latent vectors for a spatial grid (101 points), a temporal grid (103 points), and 25 distinct viscosity values. For each of the 25 viscosity values, denoted ν_k, the dataset comprises $101 \times 103 = 10,403$ latent vectors in $\mathbb{R}^{10}$. These 25 sets of vectors form the point clouds L_k for which the intrinsic dimensionality was estimated. The viscosity values ν_k span a range from 0.01 to 1.0 and were confirmed to be constant for all (x, t) points within each viscosity slice.

Initial exploratory data analysis revealed systematic changes in the distribution of latent vectors as a function of viscosity. For instance, the magnitudes of the mean and standard deviation vectors of the latent space components generally decreased as viscosity increased from 0.01 to 1.0. This suggests that the latent space representation is not static but dynamically adjusts its statistical properties in response to the physical parameter ν, indicating a potential restructuring or contraction of the latent manifold as viscosity varies.

3.2. *Intrinsic dimensionality estimates*

The intrinsic dimensionality of each of the 25 latent space point clouds L_k was estimated using the Two Nearest Neighbors (TwoNN) algorithm. This method provides a single, global estimate of the ID for a given dataset, assuming the data points lie on a manifold. The TwoNN algorithm leverages the ratio of distances to the first and second nearest neighbors to infer the local dimension, and aggregates this information to provide a global estimate.

The estimated ID values for the 25 latent spaces, corresponding to the unique viscosity values, were computed. A significant observation from these estimates is that many values substantially exceed the nominal embedding dimension of the latent space, which is 10. For example, estimated ID values reached close to 40 for intermediate viscosities. As discussed in the methods section (not provided here), ID estimators like TwoNN can yield values greater than the embedding dimension, particularly for finite datasets or complex, highly curved manifolds. Therefore, these ID estimates are interpreted not as strict geometric dimensions, but rather as quantitative measures of the *complexity*, *richness*, or *effective degrees of freedom* required to describe the structure of the learned latent manifold. A higher estimated ID indicates a more complex or space-filling arrangement of data points within the 10-dimensional embedding space. The primary focus of the analysis is on the *trend* of this complexity measure as ν changes.

3.3. *Relationship between intrinsic dimensionality and viscosity*

The core of our analysis lies in understanding how the estimated intrinsic dimensionality of the latent space varies with the viscosity parameter ν. The relationship between the estimated ID values and their corresponding viscosity values was investigated both visually and quantitatively.

Plotting the estimated ID as a function of viscosity reveals a complex and distinctly non-monotonic relationship, as shown in the top panel of Figure 1. At very low viscosities (e.g., $\nu \approx 0.01$), the estimated ID is relatively low, around 10-11, close to the embedding dimension. As viscosity increases from this low-viscosity regime towards intermediate values (up to approximately $\nu \approx 0.4$), the estimated ID shows a substantial and largely increasing trend, peaking at values near 40. Beyond this peak, in the higher viscosity regime (approximately $\nu \gtrsim 0.4$), the estimated ID exhibits a clear and significant decrease as viscosity increases further, falling to values around 24 at $\nu = 1.0$.

To quantify the overall association, Spearman's rank correlation coefficient was calculated between the sequence of viscosity values and the sequence of estimated ID values. The analysis yielded a Spearman's $\rho \approx 0.8592$ with a p-value ≈ 0.0. This strong positive correlation is statistically significant and indicates a dominant overall trend of increasing ID with viscosity across the entire range. However, as visually evident in Figure 1, this single metric does not capture the nuanced non-monotonic behavior, particularly the downturn at high viscosities.

Simple mathematical models were fitted to the ID-ν data to provide a quantitative description, acknowledging the non-monotonicity limits their descriptive power over the entire range. A linear model ($ID = a \cdot \nu + b$) yielded a slope $a \approx 22.06$, intercept $b \approx 17.80$, and an $R^2 \approx 0.313$. A logarithmic model ($ID = a \cdot \log(\nu) + b$) provided a better fit, with $a \approx 6.57$, $b \approx 37.95$, and $R^2 \approx 0.7206$. While the logarithmic model captures more of the variance, particularly the increasing phase, neither monotonic model accurately represents the ob-

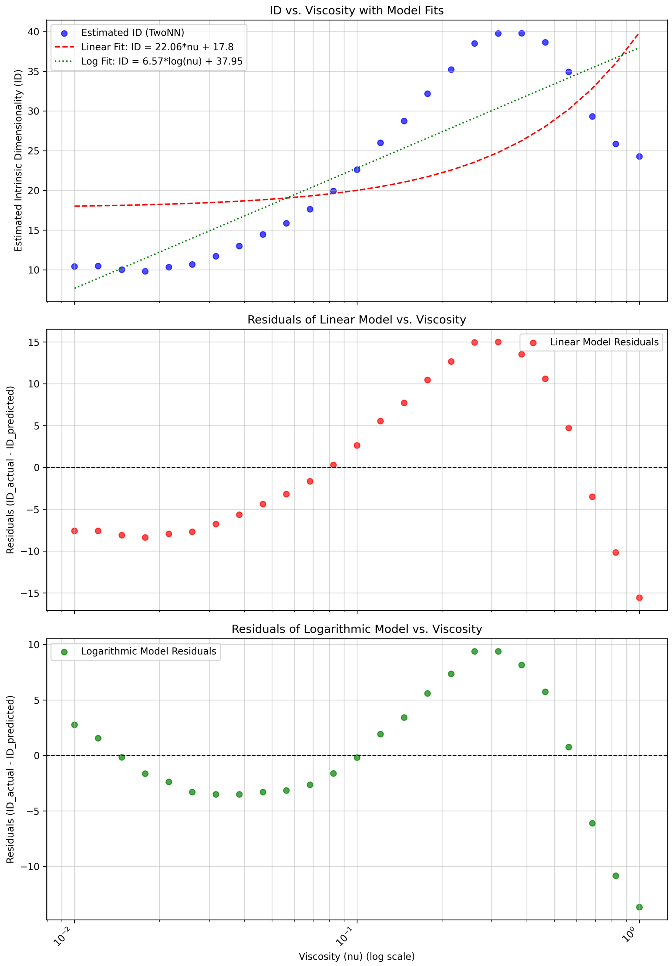

Figure 1. Estimated intrinsic dimensionality (ID) of the PINN latent space as a function of viscosity (ν). The top panel shows ID estimated by the Two Nearest Neighbors (TwoNN) algorithm versus ν on a logarithmic scale. The ID exhibits a non-monotonic trend, increasing at low ν, peaking around $\nu \approx 0.383$, and decreasing at higher ν. Simple linear and logarithmic model fits are shown, with their residuals in the middle and bottom panels respectively, illustrating their inability to capture the full non-monotonic behavior. This relationship demonstrates the dependency of latent space complexity on viscosity, suggesting the PINN's representation adapts to different physical regimes, consistent with an implicit Renormalization Group-like behavior at high ν.

served peak and subsequent decrease in ID. Residual analysis, shown in the middle and bottom panels of Figure 1, confirms that both models exhibit systematic deviations, indicating they fail to capture the full complexity of the ID-ν relationship.

3.4. *Interpretation: latent space complexity and renormalization group analogy*

The observed non-monotonic dependence of the latent space ID on viscosity provides key insights into how the PINN learns to represent physical states of varying complexity.

The physical role of viscosity (ν) in Burger's equation is that of diffusion. Higher viscosity leads to smoother solutions by damping small-scale features and counteracting the steepening effect of the non-linear term. Consequently, the physical system becomes simpler at higher ν, requiring fewer degrees of freedom for its description. Based on this, one might expect the latent space complexity, as measured by ID, to decrease monotonically with increasing viscosity.

Our results partially align with this expectation in the high-viscosity regime ($\nu \gtrsim 0.4$). As shown in Figure 1, for $\nu \gtrsim 0.4$, the estimated ID of the latent space decreases significantly as ν increases. This decrease in ID strongly suggests that the PINN learns a representation that requires fewer effective degrees of freedom to describe the physical state when the state itself is physically simpler due to high diffusion. This behavior supports our hypothesis that the PINN implicitly learns a process analogous to Renormalization Group (RG) coarse-graining. In this analogy, viscosity acts as a parameter driving the RG flow. Increasing viscosity corresponds to moving towards larger physical scales (or integrating out small scales), resulting in a simpler effective system. The corresponding decrease in latent space ID reflects the PINN's learned ability to represent these simpler states with a lower-dimensional manifold, effectively mirroring the reduction in physical degrees of freedom.

The behavior in the low-to-intermediate viscosity regime ($\nu \lesssim 0.4$), where the ID initially increases and peaks (Figure 1), adds a layer of complexity to this picture. This non-monotonicity is counter-intuitive if one only considers the simplifying effect of diffusion. Several factors could contribute to this:

- At very low viscosities, solutions might be dominated by sharp, near-shock features. While complex in gradient, the manifold of such solutions might be less diverse in structure compared to intermediate viscosities where viscous effects are sufficient to resolve these features into complex, structured viscous profiles.

- The interplay between non-linear advection and intermediate diffusion might generate a richer variety of solution patterns or transient structures than either the very low-ν (near-inviscid) or very high-ν (highly diffusive) regimes. This intermediate regime could represent a point of maximal structural complexity for the physical solutions, which the PINN then needs a more complex latent manifold to represent.

- The PINN's specific learning strategy might necessitate a more expansive use of its 10D latent space to distinguish between solutions at slightly different low-to-intermediate viscosities, where subtle changes in wave forms or viscous layers occur, leading to a higher estimated ID for these point clouds.

This initial increase suggests that the PINN's learned representation captures a transition phase where the system's complexity, from the perspective of the PINN's encoding, might temporarily increase before the strong diffusive damping at high viscosities leads to overall simplification. The peak in ID could be interpreted as occurring near a regime where the competition between non-linear steepening and viscous smoothing is most complex, potentially analogous to critical phenomena where complexity measures peak near transition points.

Overall, the systematic variation of latent space ID with viscosity demonstrates that the PINN's internal representation is intricately linked to the physical parameter governing the system's dynamics and complexity. The decrease in ID at high ν provides compelling data-driven evidence for the PINN learning an RG-like coarse-graining process. The non-monotonicity highlights that this learned relationship is not a simple one-to-one mapping of a physical parameter to a monotonically changing complexity, but rather reflects a more nuanced encoding of the physical system's behavior across different parameter regimes. These findings underscore the potential of using intrinsic dimensionality analysis of latent spaces to probe the internal workings of PINNs and connect their learned representations to fundamental physical concepts.

4. CONCLUSIONS

Understanding the internal representations learned by Physics-Informed Neural Networks (PINNs) is essential for interpreting their behavior and advancing their application in scientific computing. This study addressed the challenge of characterizing the complexity of the latent space learned by a PINN trained on the 2D Burger's equation by investigating how its intrinsic dimensionality (ID) varies with the physical parameter of viscosity, ν.

We analyzed a dataset comprising over 10,000 10-dimensional latent vectors for each of 25 distinct viscosity values, extracted from a trained PINN. Utilizing the Two Nearest Neighbors (TwoNN) algorithm, we estimated the intrinsic dimensionality of the latent space manifold for each viscosity value. The estimated ID values, interpreted as a measure of the complexity or effective degrees of freedom of the learned representation, were then analyzed as a function of viscosity.

Our analysis revealed a significant and complex non-monotonic relationship between the intrinsic dimensionality of the latent space and viscosity. While a statistically significant overall positive correlation was observed, the detailed trend showed the ID initially increasing from low viscosity values to a peak at intermediate viscosities, followed by a substantial decrease in the high-viscosity regime. For example, the estimated ID increased from approximately 10-11 at low ν to values near 40 at intermediate ν, before decreasing to around 24 at the highest ν.

These results provide empirical evidence that the PINN's learned latent space captures the underlying physical complexity of the Burger's equation across different parameter regimes. The observed decrease in latent space ID at high viscosities aligns with the physical effect of diffusion smoothing out small-scale features and simplifying the system, requiring fewer degrees of freedom for its description. This behavior supports the interpretation that the PINN implicitly learns a process analogous to Renormalization Group (RG) coarse-graining, where viscosity acts as a parameter driving a flow towards simpler representations as the physical system itself simplifies.

The non-monotonicity, particularly the initial increase in ID at low-to-intermediate viscosities, highlights the intricate nature of the learned representation. This suggests that the PINN's encoding of complexity is not a simple reflection of the physical system's degrees of freedom, but may also involve how the network distinguishes between states in regimes where the interplay of non-linear and viscous effects generates a rich variety of structures. The peak in ID could correspond to a regime of maximal representational complexity required by the network.

In conclusion, this study demonstrates that intrinsic dimensionality is a powerful quantitative tool for probing the structure and physical content of PINN latent spaces. We have shown that the complexity of the learned representation for the Burger's equation is intricately linked to the viscosity parameter, exhibiting a non-monotonic trend consistent with the physical system's behavior at high viscosity. The observed decrease in latent space ID with increasing viscosity provides compelling data-driven support for the hypothesis that PINNs can implicitly learn approximations of fundamental physical processes like RG-like coarse-graining. These findings contribute to a deeper understanding of how neural networks encode physical laws and parameters, establishing a quantitative link between learned representations and theoretical concepts in physics.